

10/537075

JC09 Rec'd PCT/PTO 01 JUN 2005

## SEQUENCE LISTING

BEST AVAILABLE COPY

<110> Kessler, Maria  
Zelinski, Thomas  
Hauer, Bernhard

<120> L-RHAMNOSE-INDUCIBLE EXPRESSION SYSTEMS

<130> 12810-00091-US

<150> PCT/EP2003/013367  
<151> 2003-11-27

<150> DE 102 56 381.0  
<151> 2002-12-02

<160> 19

<170> PatentIn version 3.3

<210> 1  
<211> 2046  
<212> DNA  
<213> Escherichia coli

<220>  
<221> misc\_feature  
<222> (288)..(1121)  
<223> coding for rhaS (positive regulator of rhaBAD operon)

<220>  
<221> misc\_feature  
<222> (1108)..(2043)  
<223> coding for rhaR (positive regulator of rhaRS operon)

<220>  
<221> protein\_bind  
<222> (56)..(72)  
<223> potential RhaS binding site

<220>  
<221> protein\_bind  
<222> (89)..(105)  
<223> potential RhaS binding site

<220>  
<221> protein\_bind  
<222> (172)..(203)  
<223> potential RhaR binding site

<220>  
<221> protein\_bind  
<222> (210)..(241)  
<223> potential RhaR binding site

<220>  
<221> misc\_feature

<222> (24)

<223> potential start of transcription (complement)

<400> 1

```

aatgtgatcc tgctgaattt cattacgacc agtctaaaaa ggcgcctgaat tcgcgacctt 60
ctcgttactg acaggaaaat gggccattgg caaccaggga aagatgaacg tgatgatgtt 120
cacaatttgc tgaattgtgg tgatgtgatg ctcaccgcat ttcctgaaaa ttcacgctgt 180
atcttgaaaa atcgacgttt ttacgtggg tttccgtcga aaatttaagg taagaacctg 240
acctcgtgat tactatttcg ccgtgttgac gacatcagga ggccagtatg accgtattac 300
atagtgtgga tttttttccg tctggtaacg cgtccgtggc gatagaacct cggctcccgc 360
aggcggattt tcctgaacat catcatgatt ttcatgaaat tgtgattgtc gaacatggca 420
cgggtattca tgtgtttaat gggcagccct ataccatcac cgggtggcacg gtctgtttcg 480
tacgcgatca tgatcggcat ctgtatgaac ataccgataa tctgtgtctg accaatgtgc 540
tgtatcgctc gccggatcga tttcagtttc tcgccgggct gaatcagttg ctgccacaag 600
agctggatgg gcagtatccg tctcactggc gcgttaacca cagcgtattg cagcagggtg 660
gacagctggg tgcacagatg gaacagcagg aaggggaaaa tgatttacct tcgaccgcca 720
gtcgcgagat cttgtttatg caattactgc tcttgctgcg taaaagcagt ttgcaggaga 780
acctgaaaaa cagcgcatca cgtctcaact tgcttctggc ctggctggag gaccattttg 840
ccgatgaggt gaattgggat gccgtggcgg atcaattttc tctttcactg cgtacgctac 900
atcggcagct taagcagcaa acgggactga cgcctcagcg atacctgaac cgcctgcgac 960
tgatgaaagc ccgacatctg ctacgccaca gcgaggccag cgttactgac atcgccatc 1020
gctgtggatt cagcgacagt aaccactttt cgacgctttt tcgccgagag ttttaactgg 1080
caccgcgtga tattcgccag ggacgggatg gctttctgca ataacgcgaa tctttctcaac 1140
gtatttgtac gccatattgc gaataatcaa cttcgttctc tggccgaggt agccacggtg 1200
gcgcacagc taaaacttct caaagatgat ttttttgcca gcgaccagca ggcagtcgct 1260
gtggctgacc gttatccgca agatgtcttt gctgaacata cacatgattt ttgtgagctg 1320
gtgattgtct ggcgcggtaa tggcctgcat gtactcaacg atcgccctta tcgcattacc 1380
cgtggcgatc tcttttacat tcatgctgac gataaacact cctacgcttc cgtaaacgat 1440
ctgggtttgc agaattattt ttattgcccg gagcgtctga agctgaatct tgactggcag 1500
ggggcgattc cgggatttaa cgccagcgca agggcaaccac actggcgctt aggtagcatg 1560
gggatggcgc aggcgcggca ggttatcggt cagcttgagc atgaaagtag tcagcatgtg 1620
ccgtttgcta acgaaatggc tgagttgctg ttcgggcagt tgggtgatgtt gctgaatcgc 1680
catcgttaca ccagtgattc gttgccgcca acatccagcg aaacgttgct ggataagctg 1740
attaccgggc tggcggctag cctgaaaagt ccctttgcgc tggataaatt ttgtgatgag 1800
gcatcgtgca gtgagcgcgt tttgcgtcag caatttcgcc agcagactgg aatgaccatc 1860
aatcaatatc tgcgacaggt cagagtgtgt catgcgcaat atcttctcca gcatagccgc 1920
ctgttaatca gtgatatttc gaccgaatgt ggctttgaag atagtaacta tttttcgggtg 1980
gtgtttacct gggaaaccgg gatgacgccc agccagtggc gtcactctca ttcgcagaaa 2040
gattaa
2046

```

<210> 2

<211> 287

<212> DNA

<213> Escherichia coli

<220>

<221> promoter

<222> (1)..(287)

<223> rhaBAD promoter fragment containing rhaS and rhaR binding sites

<400> 2

```

actggcctcc tgatgtcgtc aacacggcga aatagtaatc acgaggtcag gttcttacct 60
taaatttttc acggaaaacc acgtaaaaaa cgtcgatttt tcaagataca gcgtgaattt 120
tcaggaaatg cgggtgagcat cacatcacca caattcagca aattgtgaac atcatcacgt 180
tcatctttcc ctgggttgcca atggcccatt ttcctgtcag taacgagaag gtcgcgaatt 240
caggcgcttt ttagactggg cgtaaatgaaa ttcagcagga tcacatt
287

```

<210> 3

<211> 125  
 <212> DNA  
 <213> Escherichia coli  
 <220>  
 <221> promoter  
 <222> (1)..(125)  
 <223> rhaBAD promoter fragment containing RhaS binding site  
 <400> 3  
 ttgtgaacat catcacgttc atctttccct ggttgccaat ggcccatttt cctgtcagta 60  
 acgagaaggt cgcgaattca ggcgcttttt agactggtcg taatgaaatt cagcaggatc 120  
 acatt 125

<210> 4  
 <211> 123  
 <212> DNA  
 <213> Escherichia coli  
 <220>  
 <221> promoter  
 <222> (1)..(123)  
 <223> rhaBAD promoter fragment containing RhaS binding site  
 <400> 4  
 atcaccacaa ttcagcaaat tgtgaacatc atcacgttca tctttccctg gttgccaatg 60  
 gcccattttc ctgtcagtaa cgagaaggtc gcgaattcag gcgcttttta gactggtcgt 120  
 aat 123

<210> 5  
 <211> 51  
 <212> DNA  
 <213> Escherichia coli  
 <220>  
 <221> misc\_feature  
 <222> (1)..(51)  
 <223> palindromic RhaS binding site of rhaBAD promoter  
 <400> 5  
 atctttccct ggttgccaat ggcccatttt cctgtcagta acgagaaggt c 51

<210> 6  
 <211> 1071  
 <212> DNA  
 <213> Alcaligenes faecalis  
 <220>  
 <221> CDS  
 <222> (1)..(1068)  
 <223> coding for nitrilase  
 <400> 6  
 atg cag aca aga aaa atc gtc cgg gca gcc gcc gta cag gcc gcc tct 48  
 Met Gln Thr Arg Lys Ile Val Arg Ala Ala Ala Val Gln Ala Ala Ser  
 1 5 10 15  
 ccc aac tac gat ctg gca acg ggt gtt gat aaa acc att gag ctg gct 96  
 Pro Asn Tyr Asp Leu Ala Thr Gly Val Asp Lys Thr Ile Glu Leu Ala  
 20 25 30  
 cgt cag gcc cgc gat gag ggc tgt gac ctg atc gtg ttt ggt gaa acc 144  
 Arg Gln Ala Arg Asp Glu Gly Cys Asp Leu Ile Val Phe Gly Glu Thr  
 35 40 45

tg	g	ct	g	ccc	gga	tat	ccc	ttc	cac	gtc	tg	ct	g	ggc	gca	ccg	gcc	tg	192	
Trp	Leu	Pro	Gly	Tyr	Pro	Phe	His	Val	Trp	Leu	Gly	Ala	Pro	Ala	Trp					
50						55						60								
tc	g	ct	g	aaa	tac	agt	gcc	cgc	tac	tat	gcc	aac	tc	g	ctc	tc	g	ct	g	240
Ser	Leu	Lys	Tyr	Ser	Ala	Arg	Tyr	Tyr	Ala	Asn	Ser	Leu	Ser	Leu	Asp					
65				70				75				80								
agt	gca	gag	ttt	caa	cgc	att	gcc	cag	gcc	gca	cgg	acc	ttg	ggt	att	288				
Ser	Ala	Glu	Phe	Gln	Arg	Ile	Ala	Gln	Ala	Ala	Arg	Thr	Leu	Gly	Ile					
				85				90						95						
ttc	atc	gca	ctg	ggt	tat	agc	gag	cgc	agc	ggc	ggc	agc	ctt	tac	ctg	336				
Phe	Ile	Ala	Leu	Gly	Tyr	Ser	Glu	Arg	Ser	Gly	Gly	Ser	Leu	Tyr	Leu					
				100				105					110							
ggc	caa	tgc	ctg	atc	gac	gac	aag	ggc	gag	atg	ctg	tg	tc	g	cgt	cgc	384			
Gly	Gln	Cys	Leu	Ile	Asp	Asp	Lys	Gly	Glu	Met	Leu	Trp	Ser	Arg	Arg					
				115			120					125								
aaa	ctc	aaa	ccc	acg	cat	gta	gag	cgc	acc	gta	ttt	ggt	gaa	ggt	tat	432				
Lys	Leu	Lys	Pro	Thr	His	Val	Glu	Arg	Thr	Val	Phe	Gly	Glu	Gly	Tyr					
				130			135					140								
gcc	cgt	gat	ctg	att	gtg	tcc	gac	aca	gaa	ctg	gga	cgc	gtc	ggt	gct	480				
Ala	Arg	Asp	Leu	Ile	Val	Ser	Asp	Thr	Glu	Leu	Gly	Arg	Val	Gly	Ala					
							150					155			160					
cta	tgc	tgc	tg	gag	cat	ttg	tc	ccc	ttg	agc	aag	tac	gc	ctg	tac	528				
Leu	Cys	Cys	Trp	Glu	His	Leu	Ser	Pro	Leu	Ser	Lys	Tyr	Ala	Leu	Tyr					
				165				170						175						
tcc	cag	cat	gaa	gcc	att	cac	att	gct	gcc	tg	ccg	tc	ttt	tc	cta	576				
Ser	Gln	His	Glu	Ala	Ile	His	Ile	Ala	Ala	Trp	Pro	Ser	Phe	Ser	Leu					
				180				185					190							
tac	agc	gaa	cag	gcc	cac	gcc	ctc	agt	gcc	aag	gtg	aac	atg	gct	gcc	624				
Tyr	Ser	Glu	Gln	Ala	His	Ala	Leu	Ser	Ala	Lys	Val	Asn	Met	Ala	Ala					
				195				200					205							
tc	caa	atc	tat	tc	gtt	gaa	ggc	cag	tgc	ttt	acc	atc	gcc	gcc	agc	672				
Ser	Gln	Ile	Tyr	Ser	Val	Glu	Gly	Gln	Cys	Phe	Thr	Ile	Ala	Ala	Ser					
				210			215					220								
agt	gtg	gtc	acc	caa	gag	acg	cta	gac	atg	ctg	gaa	gtg	ggt	gaa	cac	720				
Ser	Val	Val	Thr	Gln	Glu	Thr	Leu	Asp	Met	Leu	Glu	Val	Gly	Glu	His					
				225			230				235				240					
aac	gcc	ccc	ttg	ctg	aaa	gtg	ggc	ggc	ggc	agt	tcc	atg	att	ttt	gc	768				
Asn	Ala	Pro	Leu	Leu	Lys	Val	Gly	Gly	Gly	Ser	Ser	Met	Ile	Phe	Ala					
				245				250						255						
ccg	gac	gga	cgc	aca	ctg	gct	ccc	tac	ctg	cct	cac	gat	gcc	gag	ggc	816				
Pro	Asp	Gly	Arg	Thr	Leu	Ala	Pro	Tyr	Leu	Pro	His	Asp	Ala	Glu	Gly					
				260				265					270							
ttg	atc	att	gcc	gat	ctg	aat	atg	gag	gag	att	gcc	ttc	gcc	aaa	gc	864				
Leu	Ile	Ile	Ala	Asp	Leu	Asn	Met	Glu	Glu	Ile	Ala	Phe	Ala	Lys	Ala					
				275			280						285							
atc	aat	gac	ccc	gta	ggc	cac	tat	tcc	aaa	ccc	gag	gcc	acc	cgt	ctg	912				
Ile	Asn	Asp	Pro	Val	Gly	His	Tyr	Ser	Lys	Pro	Glu	Ala	Thr	Arg	Leu					
				290			295					300								
gtg	ctg	gac	ttg	ggg	cac	cga	gac	ccc	atg	act	cgg	gtg	cac	tcc	aaa	960				
Val	Leu	Asp	Leu	Gly	His	Arg	Asp	Pro	Met	Thr	Arg	Val	His	Ser	Lys					
				305			310					315			320					
agc	gtg	acc	agg	gaa	gag	gct	ccc	gag	caa	ggt	gtg	caa	agc	aag	att	1008				
Ser	Val	Thr	Arg	Glu	Glu	Ala	Pro	Glu	Gln	Gly	Val	Gln	Ser	Lys	Ile					
				325				330						335						
gcc	tca	gtc	gct	atc	agc	cat	cca	cag	gac	tc	gac	aca	ctg	cta	gtg	1056				
Ala	Ser	Val	Ala	Ile	Ser	His	Pro	Gln	Asp	Ser	Asp	Thr	Leu	Leu	Val					
				340				345						350						

caa gag ccg tct tga  
Gln Glu Pro Ser  
355

1071

&lt;210&gt; 7

&lt;211&gt; 356

&lt;212&gt; PRT

<213> *Alcaligenes faecalis*

&lt;400&gt; 7

Met	Gln	Thr	Arg	Lys	Ile	Val	Arg	Ala	Ala	Ala	Val	Gln	Ala	Ala	Ser	1	5	10	15
Pro	Asn	Tyr	Asp	Leu	Ala	Thr	Gly	Val	Asp	Lys	Thr	Ile	Glu	Leu	Ala	20	25	30	
Arg	Gln	Ala	Arg	Asp	Glu	Gly	Cys	Asp	Leu	Ile	Val	Phe	Gly	Glu	Thr	35	40	45	
Trp	Leu	Pro	Gly	Tyr	Pro	Phe	His	Val	Trp	Leu	Gly	Ala	Pro	Ala	Trp	50	55	60	
Ser	Leu	Lys	Tyr	Ser	Ala	Arg	Tyr	Tyr	Ala	Asn	Ser	Leu	Ser	Leu	Asp	65	70	75	80
Ser	Ala	Glu	Phe	Gln	Arg	Ile	Ala	Gln	Ala	Ala	Arg	Thr	Leu	Gly	Ile	85	90	95	
Phe	Ile	Ala	Leu	Gly	Tyr	Ser	Glu	Arg	Ser	Gly	Gly	Ser	Leu	Tyr	Leu	100	105	110	
Gly	Gln	Cys	Leu	Ile	Asp	Asp	Lys	Gly	Glu	Met	Leu	Trp	Ser	Arg	Arg	115	120	125	
Lys	Leu	Lys	Pro	Thr	His	Val	Glu	Arg	Thr	Val	Phe	Gly	Glu	Gly	Tyr	130	135	140	
Ala	Arg	Asp	Leu	Ile	Val	Ser	Asp	Thr	Glu	Leu	Gly	Arg	Val	Gly	Ala	145	150	155	160
Leu	Cys	Cys	Trp	Glu	His	Leu	Ser	Pro	Leu	Ser	Lys	Tyr	Ala	Leu	Tyr	165	170	175	
Ser	Gln	His	Glu	Ala	Ile	His	Ile	Ala	Ala	Trp	Pro	Ser	Phe	Ser	Leu	180	185	190	
Tyr	Ser	Glu	Gln	Ala	His	Ala	Leu	Ser	Ala	Lys	Val	Asn	Met	Ala	Ala	195	200	205	
Ser	Gln	Ile	Tyr	Ser	Val	Glu	Gly	Gln	Cys	Phe	Thr	Ile	Ala	Ala	Ser	210	215	220	
Ser	Val	Val	Thr	Gln	Glu	Thr	Leu	Asp	Met	Leu	Glu	Val	Gly	Glu	His	225	230	235	240
Asn	Ala	Pro	Leu	Leu	Lys	Val	Gly	Gly	Gly	Ser	Ser	Met	Ile	Phe	Ala	245	250	255	
Pro	Asp	Gly	Arg	Thr	Leu	Ala	Pro	Tyr	Leu	Pro	His	Asp	Ala	Glu	Gly	260	265	270	
Leu	Ile	Ile	Ala	Asp	Leu	Asn	Met	Glu	Glu	Ile	Ala	Phe	Ala	Lys	Ala	275	280	285	
Ile	Asn	Asp	Pro	Val	Gly	His	Tyr	Ser	Lys	Pro	Glu	Ala	Thr	Arg	Leu	290	295	300	
Val	Leu	Asp	Leu	Gly	His	Arg	Asp	Pro	Met	Thr	Arg	Val	His	Ser	Lys	305	310	315	320
Ser	Val	Thr	Arg	Glu	Glu	Ala	Pro	Glu	Gln	Gly	Val	Gln	Ser	Lys	Ile	325	330	335	
Ala	Ser	Val	Ala	Ile	Ser	His	Pro	Gln	Asp	Ser	Asp	Thr	Leu	Leu	Val	340	345	350	
Gln	Glu	Pro	Ser													355			

```

<210> 8
<211> 1260
<212> DNA
<213> Escherichia coli
<220>
<221> CDS
<222> (1)..(1257)
<223> coding for rhaA (L-rhamnose isomerase)
<400> 8
atg acc act caa ctg gaa cag gcc tgg gag cta gcg aaa cag cgt ttc 48
Met Thr Thr Gln Leu Glu Gln Ala Trp Glu Leu Ala Lys Gln Arg Phe
      1              5              10              15
gcg gcg gtg ggg att gat gtc gag gag gcg ctg cgc caa ctt gat cgt 96
Ala Ala Val Gly Ile Asp Val Glu Glu Ala Leu Arg Gln Leu Asp Arg
      20              25              30
tta ccc gtt tca atg cac tgc tgg cag ggc gat gat gtt tcc ggt ttt 144
Leu Pro Val Ser Met His Cys Trp Gln Gly Asp Asp Val Ser Gly Phe
      35              40              45
gaa aac ccg gaa ggt tcg ctg acc ggg ggg att cag gcc aca ggc aat 192
Glu Asn Pro Glu Gly Ser Leu Thr Gly Gly Ile Gln Ala Thr Gly Asn
      50              55              60
tat ccg ggc aaa gcg cgt aat gcc agt gag cta cgt gcc gat ctg gaa 240
Tyr Pro Gly Lys Ala Arg Asn Ala Ser Glu Leu Arg Ala Asp Leu Glu
      65              70              75              80
cag gct atg cgg ctg att ccg ggg ccg aaa cgg ctt aat tta cat gcc 288
Gln Ala Met Arg Leu Ile Pro Gly Pro Lys Arg Leu Asn Leu His Ala
      85              90              95
atc tat ctg gaa tca gat acg cca gtc tgc cgc gac cag atc aaa cca 336
Ile Tyr Leu Glu Ser Asp Thr Pro Val Ser Arg Asp Gln Ile Lys Pro
      100             105             110
gag cac ttc aaa aac tgg gtt gaa tgg gcg aaa gcc aat cag ctc ggt 384
Glu His Phe Lys Asn Trp Val Glu Trp Ala Lys Ala Asn Gln Leu Gly
      115             120             125
ctg gat ttt aac ccc tcc tgc ttt tgc cat ccg cta agc gcc gat ggc 432
Leu Asp Phe Asn Pro Ser Cys Phe Ser His Pro Leu Ser Ala Asp Gly
      130             135             140
ttt acg ctt tcc cat gcc gac gac agc att cgc cag ttc tgg att gat 480
Phe Thr Leu Ser His Ala Asp Asp Ser Ile Arg Gln Phe Trp Ile Asp
      145             150             155             160
cac tgc aaa gcc agc cgt cgc gtt tgc gcc tat ttt ggc gag caa ctc 528
His Cys Lys Ala Ser Arg Arg Val Ser Ala Tyr Phe Gly Glu Gln Leu
      165             170             175
ggc aca cca tgc gtg atg aac atc tgg atc ccg gat ggt atg aaa gat 576
Gly Thr Pro Ser Val Met Asn Ile Trp Ile Pro Asp Gly Met Lys Asp
      180             185             190
atc acc gtt gac cgt ctc gcc ccg cgt cag cgt ctg ctg gca gca ctg 624
Ile Thr Val Asp Arg Leu Ala Pro Arg Gln Arg Leu Leu Ala Ala Leu
      195             200             205
gat gag gtg atc agc gag aag cta aac cct gcg cac cat atc gac gcc 672
Asp Glu Val Ile Ser Glu Lys Leu Asn Pro Ala His His Ile Asp Ala
      210             215             220
gtt gag agc aaa ttg ttt ggc att ggc gca gag agc tac acg gtt ggc 720
Val Glu Ser Lys Leu Phe Gly Ile Gly Ala Glu Ser Tyr Thr Val Gly
      225             230             235             240
tcc aat gag ttt tac atg ggg tat gcc acc agc cgc cag act gcg ctg 768
Ser Asn Glu Phe Tyr Met Gly Tyr Ala Thr Ser Arg Gln Thr Ala Leu

```

245	250	255	
tgc ctg gac gcc ggg cac ttc cac ccg act gaa gtg att tcc gac aag			816
Cys Leu Asp Ala Gly His Phe His Pro Thr Glu Val Ile Ser Asp Lys			
260	265	270	
att tcc gcc gcc atg ctg tat gtg ccg cag ttg ctg ctg cac gtc agc			864
Ile Ser Ala Ala Met Leu Tyr Val Pro Gln Leu Leu Leu His Val Ser			
275	280	285	
cgt ccg gtt cgc tgg gac agc gat cac gta gtg ctg ctg gat gat gaa			912
Arg Pro Val Arg Trp Asp Ser Asp His Val Val Leu Leu Asp Asp Glu			
290	295	300	
acc cag gca att gcc agt gag att gtg cgt cac gat ctg ttt gac cgg			960
Thr Gln Ala Ile Ala Ser Glu Ile Val Arg His Asp Leu Phe Asp Arg			
305	310	315	
gtg cat atc ggc ctt gac ttc ttc gat gcc tct atc aac cgc att gcc			1008
Val His Ile Gly Leu Asp Phe Phe Asp Ala Ser Ile Asn Arg Ile Ala			
325	330	335	
gcg tgg gtc att ggt aca cgc aat atg aaa aaa gcc ctg ctg cgt gcg			1056
Ala Trp Val Ile Gly Thr Arg Asn Met Lys Lys Ala Leu Leu Arg Ala			
340	345	350	
ttg ctg gaa cct acc gct gac gtg cgc aag ctg gaa gcg gcg ggc gat			1104
Leu Leu Glu Pro Thr Ala Asp Val Arg Lys Leu Glu Ala Ala Gly Asp			
355	360	365	
tac act gcg cgt ctg gca ctg ctg gaa gag cag aaa tcg ttg ccg tgg			1152
Tyr Thr Ala Arg Leu Ala Leu Leu Glu Glu Gln Lys Ser Leu Pro Trp			
370	375	380	
cag gcg gtc tgg gaa atg tat tgc caa cgt cac gat acg cca gca ggt			1200
Gln Ala Val Trp Glu Met Tyr Cys Gln Arg His Asp Thr Pro Ala Gly			
385	390	395	
agc gaa tgg ctg gag agc gtg cgg gct tat gag aaa gaa att ttg agt			1248
Ser Glu Trp Leu Glu Ser Val Arg Ala Tyr Glu Lys Glu Ile Leu Ser			
405	410	415	
cgc cgc ggg taa			1260
Arg Arg Gly			

&lt;210&gt; 9

&lt;211&gt; 419

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 9

Met Thr Thr Gln Leu Glu Gln Ala Trp Glu Leu Ala Lys Gln Arg Phe	
1 5 10 15	
Ala Ala Val Gly Ile Asp Val Glu Glu Ala Leu Arg Gln Leu Asp Arg	
20 25 30	
Leu Pro Val Ser Met His Cys Trp Gln Gly Asp Asp Val Ser Gly Phe	
35 40 45	
Glu Asn Pro Glu Gly Ser Leu Thr Gly Gly Ile Gln Ala Thr Gly Asn	
50 55 60	
Tyr Pro Gly Lys Ala Arg Asn Ala Ser Glu Leu Arg Ala Asp Leu Glu	
65 70 75 80	
Gln Ala Met Arg Leu Ile Pro Gly Pro Lys Arg Leu Asn Leu His Ala	
85 90 95	
Ile Tyr Leu Glu Ser Asp Thr Pro Val Ser Arg Asp Gln Ile Lys Pro	
100 105 110	
Glu His Phe Lys Asn Trp Val Glu Trp Ala Lys Ala Asn Gln Leu Gly	
115 120 125	
Leu Asp Phe Asn Pro Ser Cys Phe Ser His Pro Leu Ser Ala Asp Gly	

130		135		140
Phe Thr Leu Ser His	Ala Asp Asp Ser Ile Arg	Gln Phe Trp Ile Asp		
145	150	155	160	
His Cys Lys Ala Ser	Arg Arg Val Ser Ala Tyr	Phe Gly Glu Gln Leu		
	165	170	175	
Gly Thr Pro Ser Val	Met Asn Ile Trp Ile Pro	Asp Gly Met Lys Asp		
	180	185	190	
Ile Thr Val Asp Arg	Leu Ala Pro Arg Gln Arg	Leu Leu Ala Ala Leu		
	195	200	205	
Asp Glu Val Ile Ser	Glu Lys Leu Asn Pro Ala	His His Ile Asp Ala		
	210	215	220	
Val Glu Ser Lys Leu	Phe Gly Ile Gly Ala Glu	Ser Tyr Thr Val Gly		
225	230	235	240	
Ser Asn Glu Phe Tyr	Met Gly Tyr Ala Thr Ser	Arg Gln Thr Ala Leu		
	245	250	255	
Cys Leu Asp Ala Gly	His Phe His Pro Thr	Glu Val Ile Ser Asp Lys		
	260	265	270	
Ile Ser Ala Ala Met	Leu Tyr Val Pro Gln	Leu Leu Leu His Val Ser		
	275	280	285	
Arg Pro Val Arg Trp	Asp Ser Asp His Val Val	Leu Leu Asp Asp Glu		
	290	295	300	
Thr Gln Ala Ile Ala	Ser Glu Ile Val Arg His	Asp Leu Phe Asp Arg		
305	310	315	320	
Val His Ile Gly Leu	Asp Phe Phe Asp Ala Ser	Ile Asn Arg Ile Ala		
	325	330	335	
Ala Trp Val Ile Gly	Thr Arg Asn Met Lys Lys	Ala Leu Leu Arg Ala		
	340	345	350	
Leu Leu Glu Pro Thr	Ala Asp Val Arg Lys Leu	Glu Ala Ala Gly Asp		
	355	360	365	
Tyr Thr Ala Arg Leu	Ala Leu Leu Glu Glu Gln	Lys Ser Leu Pro Trp		
	370	375	380	
Gln Ala Val Trp Glu	Met Tyr Cys Gln Arg His	Asp Thr Pro Ala Gly		
385	390	395	400	
Ser Glu Trp Leu Glu	Ser Val Arg Ala Tyr Glu	Lys Glu Ile Leu Ser		
	405	410	415	
Arg Arg Gly				

&lt;210&gt; 10

&lt;211&gt; 1470

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1467)

&lt;223&gt; coding for rhaB (rhammolukinase)

&lt;400&gt; 10

atg acc ttt cgc aat tgt gtc gcc gtc gat ctc ggc gca tcc agt ggg	48
Met Thr Phe Arg Asn Cys Val Ala Val Asp Leu Gly Ala Ser Ser Gly	
1 5 10 15	
cgc gtg atg ctg gcg cgt tac gag cgt gaa tgc cgc agc ctg acg ctg	96
Arg Val Met Leu Ala Arg Tyr Glu Arg Glu Cys Arg Ser Leu Thr Leu	
20 25 30	
cgc gaa atc cat cgt ttt aac aat ggg ctg cat agt cag aac ggc tat	144
Arg Glu Ile His Arg Phe Asn Asn Gly Leu His Ser Gln Asn Gly Tyr	
35 40 45	
gtc acc tgg gat gtg gat agc ctt gaa agt gcc att cgc ctt gga tta	192



Val	Thr	Trp	Asp	Val	Asp	Ser	Leu	Glu	Ser	Ala	Ile	Arg	Leu	Gly	Leu	
50					55					60						
aac	aag	gtg	tgc	gag	gaa	ggg	att	cgt	atc	gat	agc	att	ggg	att	gat	240
Asn	Lys	Val	Cys	Glu	Glu	Gly	Ile	Arg	Ile	Asp	Ser	Ile	Gly	Ile	Asp	
65					70					75					80	
acc	tgg	ggc	gtg	gac	ttt	gtg	ctg	ctc	gac	caa	cag	ggc	cag	cgt	gtg	288
Thr	Trp	Gly	Val	Asp	Phe	Val	Leu	Leu	Asp	Gln	Gln	Gly	Gln	Arg	Val	
				85					90					95		
ggc	ctg	ccc	gtt	gct	tat	cgc	gat	agc	cgc	acc	aat	ggc	cta	atg	gcg	336
Gly	Leu	Pro	Val	Ala	Tyr	Arg	Asp	Ser	Arg	Thr	Asn	Gly	Leu	Met	Ala	
			100					105					110			
cag	gca	caa	caa	caa	ctc	ggc	aaa	cgc	gat	att	tat	caa	cgt	agc	ggc	384
Gln	Ala	Gln	Gln	Gln	Leu	Gly	Lys	Arg	Asp	Ile	Tyr	Gln	Arg	Ser	Gly	
		115					120					125				
atc	cag	ttt	ctg	ccc	ttc	aat	acg	ctt	tat	cag	ttg	cgt	gcg	ctg	acg	432
Ile	Gln	Phe	Leu	Pro	Phe	Asn	Thr	Leu	Tyr	Gln	Leu	Arg	Ala	Leu	Thr	
	130					135					140					
gag	caa	caa	cct	gaa	ctt	att	cca	cac	att	gct	cac	gct	ctg	ctg	atg	480
Glu	Gln	Gln	Pro	Glu	Leu	Ile	Pro	His	Ile	Ala	His	Ala	Leu	Leu	Met	
				145		150				155					160	
ccg	gat	tac	ttc	agt	tat	cgc	ctg	acc	ggc	aag	atg	aac	tgg	gaa	tat	528
Pro	Asp	Tyr	Phe	Ser	Tyr	Arg	Leu	Thr	Gly	Lys	Met	Asn	Trp	Glu	Tyr	
				165					170					175		
acc	aac	gcc	acg	acc	acg	caa	ctg	gtc	aat	atc	aat	agc	gac	gac	tgg	576
Thr	Asn	Ala	Thr	Thr	Thr	Gln	Leu	Val	Asn	Ile	Asn	Ser	Asp	Asp	Trp	
				180				185					190			
gac	gag	tcg	cta	ctg	gcg	tgg	agc	ggg	gcc	aac	aaa	gcc	tgg	ttt	ggc	624
Asp	Glu	Ser	Leu	Leu	Ala	Trp	Ser	Gly	Ala	Asn	Lys	Ala	Trp	Phe	Gly	
		195					200				205					
cgc	ccg	acg	cat	ccg	ggc	aat	gtc	ata	ggc	cac	tgg	att	tgc	ccg	cag	672
Arg	Pro	Thr	His	Pro	Gly	Asn	Val	Ile	Gly	His	Trp	Ile	Cys	Pro	Gln	
	210					215					220					
ggc	aac	gag	att	cca	gtg	gtc	gcc	gtt	gcc	agc	cat	gat	acc	gcc	agc	720
Gly	Asn	Glu	Ile	Pro	Val	Val	Ala	Val	Ala	Ser	His	Asp	Thr	Ala	Ser	
				225		230				235					240	
gcg	gtt	atc	gcc	tcg	ccg	tta	aac	ggc	tca	cgt	gct	gct	tat	ctc	tct	768
Ala	Val	Ile	Ala	Ser	Pro	Leu	Asn	Gly	Ser	Arg	Ala	Ala	Tyr	Leu	Ser	
				245					250					255		
tct	ggc	acc	tgg	tca	ttg	atg	ggc	ttc	gaa	agc	cag	acg	cca	ttt	acc	816
Ser	Gly	Thr	Trp	Ser	Leu	Met	Gly	Phe	Glu	Ser	Gln	Thr	Pro	Phe	Thr	
				260				265					270			
aat	gac	acg	gca	ctg	gca	gcc	aac	atc	acc	aat	gaa	ggc	ggg	gcg	gaa	864
Asn	Asp	Thr	Ala	Leu	Ala	Ala	Asn	Ile	Thr	Asn	Glu	Gly	Gly	Ala	Glu	
			275				280					285				
ggc	cgc	tat	cgg	gtg	ctg	aaa	aat	att	atg	ggc	tta	tgg	ctg	ctt	cag	912
Gly	Arg	Tyr	Arg	Val	Leu	Lys	Asn	Ile	Met	Gly	Leu	Trp	Leu	Leu	Gln	
	290					295				300						
cga	gtg	ctt	cag	gag	cag	caa	atc	aac	gat	ctt	ccg	gcg	ctt	atc	tcc	960
Arg	Val	Leu	Gln	Glu	Gln	Gln	Ile	Asn	Asp	Leu	Pro	Ala	Leu	Ile	Ser	
	305				310					315					320	
gcg	aca	cag	gca	ctt	ccg	gct	tgc	cgc	ttc	att	atc	aat	ccc	aat	gac	1008
Ala	Thr	Gln	Ala	Leu	Pro	Ala	Cys	Arg	Phe	Ile	Ile	Asn	Pro	Asn	Asp	
				325					330					335		
gat	cgc	ttt	att	aat	cct	gag	acg	atg	tgc	agc	gaa	att	cag	gct	gcg	1056
Asp	Arg	Phe	Ile	Asn	Pro	Glu	Thr	Met	Cys	Ser	Glu	Ile	Gln	Ala	Ala	
			340					345					350			
tgt	cgg	gaa	acg	gcg	caa	ccg	atc	ccg	gaa	agt	gat	gct	gaa	ctg	gcg	1104

Cys	Arg	Glu	Thr	Ala	Gln	Pro	Ile	Pro	Glu	Ser	Asp	Ala	Glu	Leu	Ala		
		355					360					365					
cgc	tgc	att	ttc	gac	agt	ctg	gcg	ctg	ctg	tat	gcc	gat	gtg	ttg	cat	1152	
Arg	Cys	Ile	Phe	Asp	Ser	Leu	Ala	Leu	Leu	Tyr	Ala	Asp	Val	Leu	His		
		370					375				380						
gag	ctg	gcg	cag	ctg	cgc	ggg	gaa	gat	ttc	tcg	caa	ctg	cat	att	gtc	1200	
Glu	Leu	Ala	Gln	Leu	Arg	Gly	Glu	Asp	Phe	Ser	Gln	Leu	His	Ile	Val		
		385				390				395					400		
ggc	gga	ggc	tgc	cag	aac	acg	ctg	ctc	aac	cag	cta	tgc	gcc	gat	gcc	1248	
Gly	Gly	Gly	Cys	Gln	Asn	Thr	Leu	Leu	Asn	Gln	Leu	Cys	Ala	Asp	Ala		
				405					410					415			
tgc	ggg	att	cgg	gtg	atc	gcc	ggg	cct	gtt	gaa	gcc	tcg	acg	ctc	ggc	1296	
Cys	Gly	Ile	Arg	Val	Ile	Ala	Gly	Pro	Val	Glu	Ala	Ser	Thr	Leu	Gly		
			420					425					430				
aat	atc	ggc	atc	cag	tta	atg	acg	ctg	gat	gaa	ctc	aac	aat	gtg	gat	1344	
Asn	Ile	Gly	Ile	Gln	Leu	Met	Thr	Leu	Asp	Glu	Leu	Asn	Asn	Val	Asp		
		435						440				445					
gat	ttc	cgt	cag	gtc	gtc	agc	acc	acc	gcg	aat	ctg	acc	acc	ttt	acc	1392	
Asp	Phe	Arg	Gln	Val	Val	Ser	Thr	Thr	Ala	Asn	Leu	Thr	Thr	Phe	Thr		
		450				455				460							
cct	aat	cct	gac	agt	gaa	att	gcc	cac	tat	gtg	gcg	cag	att	cac	tct	1440	
Pro	Asn	Pro	Asp	Ser	Glu	Ile	Ala	His	Tyr	Val	Ala	Gln	Ile	His	Ser		
		465			470					475					480		
aca	cga	cag	aca	aag	gag	ctt	tgc	gca	tga							1470	
Thr	Arg	Gln	Thr	Lys	Glu	Leu	Cys	Ala									
				485													

&lt;210&gt; 11

&lt;211&gt; 489

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 11

Met	Thr	Phe	Arg	Asn	Cys	Val	Ala	Val	Asp	Leu	Gly	Ala	Ser	Ser	Gly		
1				5					10					15			
Arg	Val	Met	Leu	Ala	Arg	Tyr	Glu	Arg	Glu	Cys	Arg	Ser	Leu	Thr	Leu		
			20					25					30				
Arg	Glu	Ile	His	Arg	Phe	Asn	Asn	Gly	Leu	His	Ser	Gln	Asn	Gly	Tyr		
			35				40					45					
Val	Thr	Trp	Asp	Val	Asp	Ser	Leu	Glu	Ser	Ala	Ile	Arg	Leu	Gly	Leu		
	50				55					60							
Asn	Lys	Val	Cys	Glu	Gly	Ile	Arg	Ile	Asp	Ser	Ile	Gly	Ile	Asp			
65				70					75					80			
Thr	Trp	Gly	Val	Asp	Phe	Val	Leu	Leu	Asp	Gln	Gln	Gly	Gln	Arg	Val		
			85						90				95				
Gly	Leu	Pro	Val	Ala	Tyr	Arg	Asp	Ser	Arg	Thr	Asn	Gly	Leu	Met	Ala		
			100				105					110					
Gln	Ala	Gln	Gln	Gln	Leu	Gly	Lys	Arg	Asp	Ile	Tyr	Gln	Arg	Ser	Gly		
			115				120					125					
Ile	Gln	Phe	Leu	Pro	Phe	Asn	Thr	Leu	Tyr	Gln	Leu	Arg	Ala	Leu	Thr		
			130			135					140						
Glu	Gln	Gln	Pro	Glu	Leu	Ile	Pro	His	Ile	Ala	His	Ala	Leu	Leu	Met		
145				150					155						160		
Pro	Asp	Tyr	Phe	Ser	Tyr	Arg	Leu	Thr	Gly	Lys	Met	Asn	Trp	Glu	Tyr		
			165					170					175				
Thr	Asn	Ala	Thr	Thr	Thr	Gln	Leu	Val	Asn	Ile	Asn	Ser	Asp	Asp	Trp		
			180					185					190				

Asp Glu Ser Leu Leu Ala Trp Ser Gly Ala Asn Lys Ala Trp Phe Gly  
           195                          200                          205  
 Arg Pro Thr His Pro Gly Asn Val Ile Gly His Trp Ile Cys Pro Gln  
           210                          215                          220  
 Gly Asn Glu Ile Pro Val Val Ala Val Ala Ser His Asp Thr Ala Ser  
 225                          230                          235                          240  
 Ala Val Ile Ala Ser Pro Leu Asn Gly Ser Arg Ala Ala Tyr Leu Ser  
                           245                          250                          255  
 Ser Gly Thr Trp Ser Leu Met Gly Phe Glu Ser Gln Thr Pro Phe Thr  
                           260                          265                          270  
 Asn Asp Thr Ala Leu Ala Ala Asn Ile Thr Asn Glu Gly Gly Ala Glu  
                           275                          280                          285  
 Gly Arg Tyr Arg Val Leu Lys Asn Ile Met Gly Leu Trp Leu Leu Gln  
                           290                          295                          300  
 Arg Val Leu Gln Glu Gln Gln Ile Asn Asp Leu Pro Ala Leu Ile Ser  
 305                          310                          315                          320  
 Ala Thr Gln Ala Leu Pro Ala Cys Arg Phe Ile Ile Asn Pro Asn Asp  
                           325                          330                          335  
 Asp Arg Phe Ile Asn Pro Glu Thr Met Cys Ser Glu Ile Gln Ala Ala  
                           340                          345                          350  
 Cys Arg Glu Thr Ala Gln Pro Ile Pro Glu Ser Asp Ala Glu Leu Ala  
                           355                          360                          365  
 Arg Cys Ile Phe Asp Ser Leu Ala Leu Leu Tyr Ala Asp Val Leu His  
                           370                          375                          380  
 Glu Leu Ala Gln Leu Arg Gly Glu Asp Phe Ser Gln Leu His Ile Val  
 385                          390                          395                          400  
 Gly Gly Gly Cys Gln Asn Thr Leu Leu Asn Gln Leu Cys Ala Asp Ala  
                           405                          410                          415  
 Cys Gly Ile Arg Val Ile Ala Gly Pro Val Glu Ala Ser Thr Leu Gly  
                           420                          425                          430  
 Asn Ile Gly Ile Gln Leu Met Thr Leu Asp Glu Leu Asn Asn Val Asp  
                           435                          440                          445  
 Asp Phe Arg Gln Val Val Ser Thr Thr Ala Asn Leu Thr Thr Phe Thr  
                           450                          455                          460  
 Pro Asn Pro Asp Ser Glu Ile Ala His Tyr Val Ala Gln Ile His Ser  
 465                          470                          475                          480  
 Thr Arg Gln Thr Lys Glu Leu Cys Ala  
                           485

<210> 12

<211> 825

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(822)

<223> coding for rhaD (rhamnulose-phosphate aldolase)

<400> 12

atg	caa	aac	att	act	cag	tcc	tgg	ttt	gtc	cag	gga	atg	atc	aaa	gcc	48
Met	Gln	Asn	Ile	Thr	Gln	Ser	Trp	Phe	Val	Gln	Gly	Met	Ile	Lys	Ala	
1				5				10						15		
acc	acc	gac	gcc	tgg	ctg	aaa	ggc	tgg	gat	gag	cgc	aac	ggc	ggc	aac	96
Thr	Thr	Asp	Ala	Trp	Leu	Lys	Gly	Trp	Asp	Glu	Arg	Asn	Gly	Gly	Asn	
			20					25					30			
ctg	acg	cta	cgc	ctg	gat	gac	gcc	gat	atc	gca	cca	tat	cac	gac	aat	144
Leu	Thr	Leu	Arg	Leu	Asp	Asp	Ala	Asp	Ile	Ala	Pro	Tyr	His	Asp	Asn	

			35			40						45						
ttc	cac	caa	caa	ccg	cgc	tat	atc	ccg	ctc	agc	cag	ccc	atg	cct	tta	192		
Phe	His	Gln	Gln	Pro	Arg	Tyr	Ile	Pro	Leu	Ser	Gln	Pro	Met	Pro	Leu			
			50				55				60							
ctg	gca	aat	aca	ccg	ttt	att	gtc	acc	ggc	tcg	ggc	aaa	ttc	ttc	cgt	240		
Leu	Ala	Asn	Thr	Pro	Phe	Ile	Val	Thr	Gly	Ser	Gly	Lys	Phe	Phe	Arg			
			65				70				75				80			
aac	gtc	cag	ctt	gat	cct	gcg	gct	aac	tta	ggc	atc	gta	aaa	gtc	gac	288		
Asn	Val	Gln	Leu	Asp	Pro	Ala	Ala	Asn	Leu	Gly	Ile	Val	Lys	Val	Asp			
			85			90			95									
agc	gac	ggc	gcg	ggc	tac	cac	att	ctt	tgg	ggg	tta	acc	aac	gaa	gcc	336		
Ser	Asp	Gly	Ala	Gly	Tyr	His	Ile	Leu	Trp	Gly	Leu	Thr	Asn	Glu	Ala			
			100			105			110									
gtc	ccc	act	tcc	gaa	ctt	ccg	gct	cac	ttc	ctt	tcc	cac	tgc	gag	cgc	384		
Val	Pro	Thr	Ser	Glu	Leu	Pro	Ala	His	Phe	Leu	Ser	His	Cys	Glu	Arg			
			115			120			125									
att	aaa	gcc	acc	aac	ggc	aaa	gat	cgg	gtg	atc	atg	cac	tgc	cac	gcc	432		
Ile	Lys	Ala	Thr	Asn	Gly	Lys	Asp	Arg	Val	Ile	Met	His	Cys	His	Ala			
			130			135			140									
acc	aac	ctg	atc	gcc	ctc	acc	tat	gta	ctt	gaa	aac	gac	acc	gcg	gtc	480		
Thr	Asn	Leu	Ile	Ala	Leu	Thr	Tyr	Val	Leu	Glu	Asn	Asp	Thr	Ala	Val			
			145			150			155			160						
ttc	act	cgc	caa	ctg	tgg	gaa	ggc	agc	acc	gag	tgt	ctg	gtg	gta	ttc	528		
Phe	Thr	Arg	Gln	Leu	Trp	Glu	Gly	Ser	Thr	Glu	Cys	Leu	Val	Val	Phe			
			165			170			175									
ccg	gat	ggc	gtt	ggc	att	ttg	ccg	tgg	atg	gtg	ccc	ggc	acg	gac	gaa	576		
Pro	Asp	Gly	Val	Gly	Ile	Leu	Pro	Trp	Met	Val	Pro	Gly	Thr	Asp	Glu			
			180			185			190									
atc	ggc	cag	gcg	acc	gca	caa	gag	atg	caa	aaa	cat	tcg	ctg	gtg	ttg	624		
Ile	Gly	Gln	Ala	Thr	Ala	Gln	Glu	Met	Gln	Lys	His	Ser	Leu	Val	Leu			
			195			200			205									
tgg	ccc	ttc	cac	ggc	gtc	ttc	ggc	agc	gga	ccg	acg	ctg	gat	gaa	acc	672		
Trp	Pro	Phe	His	Gly	Val	Phe	Gly	Ser	Gly	Pro	Thr	Leu	Asp	Glu	Thr			
			210			215			220									
ttc	ggg	tta	atc	gac	acc	gca	gaa	aaa	tca	gca	caa	gta	tta	gtg	aag	720		
Phe	Gly	Leu	Ile	Asp	Thr	Ala	Glu	Lys	Ser	Ala	Gln	Val	Leu	Val	Lys			
			225			230			235			240						
gtt	tat	tcg	atg	ggc	ggc	atg	aaa	cag	acc	atc	agc	cgt	gaa	gag	ttg	768		
Val	Tyr	Ser	Met	Gly	Gly	Met	Lys	Gln	Thr	Ile	Ser	Arg	Glu	Glu	Leu			
			245			250			255									
ata	gcg	ctc	ggc	aag	cgt	ttc	ggc	gtt	acg	cca	ctc	gcc	agt	gcg	ctg	816		
Ile	Ala	Leu	Gly	Lys	Arg	Phe	Gly	Val	Thr	Pro	Leu	Ala	Ser	Ala	Leu			
			260			265			270									
gcg	ctg	taa														825		
Ala	Leu																	

<210> 13

<211> 274

<212> PRT

<213> Escherichia coli

<400> 13

Met Gln Asn Ile Thr Gln Ser Trp Phe Val Gln Gly Met Ile Lys Ala

**1**

5

1

11

Thr Thr Asp Ala Trp Leu Lys Gly Trp Asp Glu Arg Asn Gly Gly Asn

2

2

3

Leu Thr Leu Arg Leu Asp Asp Ala Asp Ile Ala Pro Tyr His Asp Asn

[illegible]

```

<210> 14
<211> 939
<212> DNA
<213> Escherichia coli
<220>
<221> CDS
<222> (1)..(936)
<223> coding for rhaR (positive regulator for rhaRS operon)
<400> 14
atg gct ttc tgc aat aac gcg aat ctt ctc aac gta ttt gta cgc cat 48
Met Ala Phe Cys Asn Asn Ala Asn Leu Leu Asn Val Phe Val Arg His
      1              5              10              15
att gcg aat aat caa ctt cgt tct ctg gcc gag gta gcc acg gtg gcg 96
Ile Ala Asn Asn Gln Leu Arg Ser Leu Ala Glu Val Ala Thr Val Ala
      20              25              30
cat cag tta aaa ctt ctc aaa gat gat ttt ttt gcc agc gac cag cag 144
His Gln Leu Lys Leu Leu Lys Asp Asp Phe Phe Ala Ser Asp Gln Gln
      35              40              45
gca gtc gct gtg gct gac cgt tat ccg caa gat gtc ttt gct gaa cat 192
Ala Val Ala Val Ala Asp Arg Tyr Pro Gln Asp Val Phe Ala Glu His
      50              55              60
aca cat gat ttt tgt gag ctg gtg att gtc tgg cgc ggt aat ggc ctg 240
Thr His Asp Phe Cys Glu Leu Val Ile Val Trp Arg Gly Asn Gly Leu
      65              70              75              80
cat gta ctc aac gat cgc cct tat cgc att acc cgt ggc gat ctc ttt 288

```

His	Val	Leu	Asn	Asp	Arg	Pro	Tyr	Arg	Ile	Thr	Arg	Gly	Asp	Leu	Phe		
				85					90					95			
tac	att	cat	gct	gac	gat	aaa	cac	tcc	tac	gct	tcc	gtt	aac	gat	ctg	336	
Tyr	Ile	His	Ala	Asp	Asp	Lys	His	Ser	Tyr	Ala	Ser	Val	Asn	Asp	Leu		
			100					105					110				
gtt	ttg	cag	aat	att	att	tat	tgc	ccg	gag	cgt	ctg	aag	ctg	aat	ctt	384	
Val	Leu	Gln	Asn	Ile	Ile	Tyr	Cys	Pro	Glu	Arg	Leu	Lys	Leu	Asn	Leu		
		115					120					125					
gac	tgg	cag	ggg	gcg	att	ccg	gga	ttt	aac	gcc	agc	gca	ggg	caa	cca	432	
Asp	Trp	Gln	Gly	Ala	Ile	Pro	Gly	Phe	Asn	Ala	Ser	Ala	Gly	Gln	Pro		
		130				135					140						
cac	tgg	cgc	tta	ggt	agc	atg	ggg	atg	gcg	cag	gcg	cgg	cag	gtt	atc	480	
His	Trp	Arg	Leu	Gly	Ser	Met	Gly	Met	Ala	Gln	Ala	Arg	Gln	Val	Ile		
145					150				155					160			
ggt	cag	ctt	gag	cat	gaa	agt	agt	cag	cat	gtg	ccg	ttt	gct	aac	gaa	528	
Gly	Gln	Leu	Glu	His	Glu	Ser	Ser	Gln	His	Val	Pro	Phe	Ala	Asn	Glu		
			165					170				175					
atg	gct	gag	ttg	ctg	ttc	ggg	cag	ttg	gtg	atg	ttg	ctg	aat	cgc	cat	576	
Met	Ala	Glu	Leu	Leu	Phe	Gly	Gln	Leu	Val	Met	Leu	Leu	Asn	Arg	His		
			180					185				190					
cgt	tac	acc	agt	gat	tcg	ttg	ccg	cca	aca	tcc	agc	gaa	acg	ttg	ctg	624	
Arg	Tyr	Thr	Ser	Asp	Ser	Leu	Pro	Pro	Thr	Ser	Ser	Glu	Thr	Leu	Leu		
		195				200						205					
gat	aag	ctg	att	acc	cgg	ctg	gcg	gct	agc	ctg	aaa	agt	ccc	ttt	gcg	672	
Asp	Lys	Leu	Ile	Thr	Arg	Leu	Ala	Ala	Ser	Leu	Lys	Ser	Pro	Phe	Ala		
	210				215						220						
ctg	gat	aaa	ttt	tgt	gat	gag	gca	tcg	tgc	agt	gag	cgc	gtt	ttg	cgt	720	
Leu	Asp	Lys	Phe	Cys	Asp	Glu	Ala	Ser	Cys	Ser	Glu	Arg	Val	Leu	Arg		
225				230					235				240				
cag	caa	ttt	cgc	cag	cag	act	gga	atg	acc	atc	aat	caa	tat	ctg	cga	768	
Gln	Gln	Phe	Arg	Gln	Gln	Thr	Gly	Met	Thr	Ile	Asn	Gln	Tyr	Leu	Arg		
		245						250				255					
cag	gtc	aga	gtg	tgt	cat	gcg	caa	tat	ctt	ctc	cag	cat	agc	cgc	ctg	816	
Gln	Val	Arg	Val	Cys	His	Ala	Gln	Tyr	Leu	Leu	Gln	His	Ser	Arg	Leu		
		260				265					270						
tta	atc	agt	gat	att	tcg	acc	gaa	tgt	ggc	ttt	gaa	gat	agt	aac	tat	864	
Leu	Ile	Ser	Asp	Ile	Ser	Thr	Glu	Cys	Gly	Phe	Glu	Asp	Ser	Asn	Tyr		
		275				280					285						
ttt	tcg	gtg	gtg	ttt	acc	cgg	gaa	acc	ggg	atg	acg	ccc	agc	cag	tgg	912	
Phe	Ser	Val	Val	Phe	Thr	Arg	Glu	Thr	Gly	Met	Thr	Pro	Ser	Gln	Trp		
	290				295						300						
cgt	cat	ctc	aat	tcg	cag	aaa	gat	taa								939	
Arg	His	Leu	Asn	Ser	Gln	Lys	Asp										
305					310												

&lt;210&gt; 15

&lt;211&gt; 312

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 15

Met	Ala	Phe	Cys	Asn	Asn	Ala	Asn	Leu	Leu	Asn	Val	Phe	Val	Arg	His		
1				5				10						15			
Ile	Ala	Asn	Asn	Gln	Leu	Arg	Ser	Leu	Ala	Glu	Val	Ala	Thr	Val	Ala		
		20					25					30					
His	Gln	Leu	Lys	Leu	Leu	Lys	Asp	Asp	Phe	Phe	Ala	Ser	Asp	Gln	Gln		
	35						40					45					

Ala	Val	Ala	Val	Ala	Asp	Arg	Tyr	Pro	Gln	Asp	Val	Phe	Ala	Glu	His
50						55					60				
Thr	His	Asp	Phe	Cys	Glu	Leu	Val	Ile	Val	Trp	Arg	Gly	Asn	Gly	Leu
65					70					75					80
His	Val	Leu	Asn	Asp	Arg	Pro	Tyr	Arg	Ile	Thr	Arg	Gly	Asp	Leu	Phe
			85						90					95	
Tyr	Ile	His	Ala	Asp	Asp	Lys	His	Ser	Tyr	Ala	Ser	Val	Asn	Asp	Leu
			100					105					110		
Val	Leu	Gln	Asn	Ile	Ile	Tyr	Cys	Pro	Glu	Arg	Leu	Lys	Leu	Asn	Leu
		115					120					125			
Asp	Trp	Gln	Gly	Ala	Ile	Pro	Gly	Phe	Asn	Ala	Ser	Ala	Gly	Gln	Pro
130						135					140				
His	Trp	Arg	Leu	Gly	Ser	Met	Gly	Met	Ala	Gln	Ala	Arg	Gln	Val	Ile
145					150					155					160
Gly	Gln	Leu	Glu	His	Glu	Ser	Ser	Gln	His	Val	Pro	Phe	Ala	Asn	Glu
				165					170					175	
Met	Ala	Glu	Leu	Leu	Phe	Gly	Gln	Leu	Val	Met	Leu	Leu	Asn	Arg	His
			180					185					190		
Arg	Tyr	Thr	Ser	Asp	Ser	Leu	Pro	Pro	Thr	Ser	Ser	Glu	Thr	Leu	Leu
		195				200						205			
Asp	Lys	Leu	Ile	Thr	Arg	Leu	Ala	Ala	Ser	Leu	Lys	Ser	Pro	Phe	Ala
210						215					220				
Leu	Asp	Lys	Phe	Cys	Asp	Glu	Ala	Ser	Cys	Ser	Glu	Arg	Val	Leu	Arg
225					230					235					240
Gln	Gln	Phe	Arg	Gln	Gln	Thr	Gly	Met	Thr	Ile	Asn	Gln	Tyr	Leu	Arg
				245					250					255	
Gln	Val	Arg	Val	Cys	His	Ala	Gln	Tyr	Leu	Leu	Gln	His	Ser	Arg	Leu
			260					265					270		
Leu	Ile	Ser	Asp	Ile	Ser	Thr	Glu	Cys	Gly	Phe	Glu	Asp	Ser	Asn	Tyr
		275					280					285			
Phe	Ser	Val	Val	Phe	Thr	Arg	Glu	Thr	Gly	Met	Thr	Pro	Ser	Gln	Trp
290						295					300				
Arg	His	Leu	Asn	Ser	Gln	Lys	Asp								
305					310										

&lt;210&gt; 16

&lt;211&gt; 837

&lt;212&gt; DNA

&lt;213&gt; Escherichia coli

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(834)

&lt;223&gt; coding for rhaS (positive regulator of rhaBAD operon)

&lt;400&gt; 16

atg	acc	gta	tta	cat	agt	gtg	gat	ttt	ttt	ccg	tct	ggt	aac	gcg	tcc	48
Met	Thr	Val	Leu	His	Ser	Val	Asp	Phe	Phe	Pro	Ser	Gly	Asn	Ala	Ser	
1				5				10					15			
gtg	gcg	ata	gaa	ccc	cgg	ctc	ccg	cag	gcg	gat	ttt	cct	gaa	cat	cat	96
Val	Ala	Ile	Glu	Pro	Arg	Leu	Pro	Gln	Ala	Asp	Phe	Pro	Glu	His	His	
			20					25				30				
cat	gat	ttt	cat	gaa	att	gtg	att	gtc	gaa	cat	ggc	acg	ggt	att	cat	144
His	Asp	Phe	His	Glu	Ile	Val	Ile	Val	Glu	His	Gly	Thr	Gly	Ile	His	
		35				40					45					
gtg	ttt	aat	ggg	cag	ccc	tat	acc	atc	acc	ggt	ggc	acg	gtc	tgt	ttc	192
Val	Phe	Asn	Gly	Gln	Pro	Tyr	Thr	Ile	Thr	Gly	Gly	Thr	Val	Cys	Phe	
50						55					60					

gta	cgc	gat	cat	gat	cgg	cat	ctg	tat	gaa	cat	acc	gat	aat	ctg	tgt	240
Val	Arg	Asp	His	Asp	Arg	His	Leu	Tyr	Glu	His	Thr	Asp	Asn	Leu	Cys	
65					70				75					80		
ctg	acc	aat	gtg	ctg	tat	cgc	tcg	ccg	gat	cga	ttt	cag	ttt	ctc	gcc	288
Leu	Thr	Asn	Val	Leu	Tyr	Arg	Ser	Pro	Asp	Arg	Phe	Gln	Phe	Leu	Ala	
			85					90						95		
ggg	ctg	aat	cag	ttg	ctg	cca	caa	gag	ctg	gat	ggg	cag	tat	ccg	tct	336
Gly	Leu	Asn	Gln	Leu	Leu	Pro	Gln	Glu	Leu	Asp	Gly	Gln	Tyr	Pro	Ser	
			100					105					110			
cac	tgg	cgc	gtt	aac	cac	agc	gta	ttg	cag	cag	gtg	cga	cag	ctg	gtt	384
His	Trp	Arg	Val	Asn	His	Ser	Val	Leu	Gln	Gln	Val	Arg	Gln	Leu	Val	
			115				120					125				
gca	cag	atg	gaa	cag	cag	gaa	ggg	gaa	aat	gat	tta	ccc	tcg	acc	gcc	432
Ala	Gln	Met	Glu	Gln	Gln	Glu	Gly	Glu	Asn	Asp	Leu	Pro	Ser	Thr	Ala	
			130			135					140					
agt	cgc	gag	atc	ttg	ttt	atg	caa	tta	ctg	ctc	ttg	ctg	cgt	aaa	agc	480
Ser	Arg	Glu	Ile	Leu	Phe	Met	Gln	Leu	Leu	Leu	Leu	Leu	Arg	Lys	Ser	
					150				155					160		
agt	ttg	cag	gag	aac	ctg	gaa	aac	agc	gca	tca	cgt	ctc	aac	ttg	ctt	528
Ser	Leu	Gln	Glu	Asn	Leu	Glu	Asn	Ser	Ala	Ser	Arg	Leu	Asn	Leu	Leu	
				165				170					175			
ctg	gcc	tgg	ctg	gag	gac	cat	ttt	gcc	gat	gag	gtg	aat	tgg	gat	gcc	576
Leu	Ala	Trp	Leu	Glu	Asp	His	Phe	Ala	Asp	Glu	Val	Asn	Trp	Asp	Ala	
			180					185					190			
gtg	gcg	gat	caa	ttt	tct	ctt	tca	ctg	cgt	acg	cta	cat	cgg	cag	ctt	624
Val	Ala	Asp	Gln	Phe	Ser	Leu	Ser	Leu	Arg	Thr	Leu	His	Arg	Gln	Leu	
			195				200					205				
aag	cag	caa	acg	gga	ctg	acg	cct	cag	cga	tac	ctg	aac	cgc	ctg	cga	672
Lys	Gln	Gln	Thr	Gly	Leu	Thr	Pro	Gln	Arg	Tyr	Leu	Asn	Arg	Leu	Arg	
			210			215				220						
ctg	atg	aaa	gcc	cga	cat	ctg	cta	cgc	cac	agc	gag	gcc	agc	gtt	act	720
Leu	Met	Lys	Ala	Arg	His	Leu	Leu	Arg	His	Ser	Glu	Ala	Ser	Val	Thr	
					230				235					240		
gac	atc	gcc	tat	cgc	tgt	gga	ttc	agc	gac	agt	aac	cac	ttt	tcg	acg	768
Asp	Ile	Ala	Tyr	Arg	Cys	Gly	Phe	Ser	Asp	Ser	Asn	His	Phe	Ser	Thr	
				245				250					255			
ctt	ttt	cgc	cga	gag	ttt	aac	tgg	tca	ccg	cgt	gat	att	cgc	cag	gga	816
Leu	Phe	Arg	Arg	Glu	Phe	Asn	Trp	Ser	Pro	Arg	Asp	Ile	Arg	Gln	Gly	
			260					265				270				
cgg	gat	ggc	ttt	ctg	caa	taa										837
Arg	Asp	Gly	Phe	Leu	Gln											
			275													

&lt;210&gt; 17

&lt;211&gt; 278

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 17

Met	Thr	Val	Leu	His	Ser	Val	Asp	Phe	Phe	Pro	Ser	Gly	Asn	Ala	Ser	
1				5					10					15		
Val	Ala	Ile	Glu	Pro	Arg	Leu	Pro	Gln	Ala	Asp	Phe	Pro	Glu	His	His	
			20					25					30			
His	Asp	Phe	His	Glu	Ile	Val	Ile	Val	Glu	His	Gly	Thr	Gly	Ile	His	
		35					40					45				
Val	Phe	Asn	Gly	Gln	Pro	Tyr	Thr	Ile	Thr	Gly	Gly	Thr	Val	Cys	Phe	
		50					55				60					



Val Arg Asp His Asp Arg His Leu Tyr Glu His Thr Asp Asn Leu Cys  
 65 70 75 80  
 Leu Thr Asn Val Leu Tyr Arg Ser Pro Asp Arg Phe Gln Phe Leu Ala  
 85 90 95  
 Gly Leu Asn Gln Leu Leu Pro Gln Glu Leu Asp Gly Gln Tyr Pro Ser  
 100 105 110  
 His Trp Arg Val Asn His Ser Val Leu Gln Gln Val Arg Gln Leu Val  
 115 120 125  
 Ala Gln Met Glu Gln Gln Glu Gly Glu Asn Asp Leu Pro Ser Thr Ala  
 130 135 140  
 Ser Arg Glu Ile Leu Phe Met Gln Leu Leu Leu Leu Arg Lys Ser  
 145 150 155 160  
 Ser Leu Gln Glu Asn Leu Glu Asn Ser Ala Ser Arg Leu Asn Leu Leu  
 165 170 175  
 Leu Ala Trp Leu Glu Asp His Phe Ala Asp Glu Val Asn Trp Asp Ala  
 180 185 190  
 Val Ala Asp Gln Phe Ser Leu Ser Leu Arg Thr Leu His Arg Gln Leu  
 195 200 205  
 Lys Gln Gln Thr Gly Leu Thr Pro Gln Arg Tyr Leu Asn Arg Leu Arg  
 210 215 220  
 Leu Met Lys Ala Arg His Leu Leu Arg His Ser Glu Ala Ser Val Thr  
 225 230 235 240  
 Asp Ile Ala Tyr Arg Cys Gly Phe Ser Asp Ser Asn His Phe Ser Thr  
 245 250 255  
 Leu Phe Arg Arg Glu Phe Asn Trp Ser Pro Arg Asp Ile Arg Gln Gly  
 260 265 270  
 Arg Asp Gly Phe Leu Gln  
 275

<210> 18

<211> 1035

<212> DNA

<213> Escherichia coli

<220>

<221> CDS

<222> (1)..(1032)

<223> coding for rhaT (rhamnose transport protein)

<400> 18

atg agt aac gcg att acg atg ggg ata ttt tgg cat ttg atc ggc gcg 48  
 Met Ser Asn Ala Ile Thr Met Gly Ile Phe Trp His Leu Ile Gly Ala  
 1 5 10 15  
 gcc agt gca gcc tgt ttt tac gct ccg ttc aaa aaa gta aaa aaa tgg 96  
 Ala Ser Ala Ala Cys Phe Tyr Ala Pro Phe Lys Lys Val Lys Lys Trp  
 20 25 30  
 tca tgg gaa acc atg tgg tca gtc ggt ggg att gtt tcg tgg att att 144  
 Ser Trp Glu Thr Met Trp Ser Val Gly Gly Ile Val Ser Trp Ile Ile  
 35 40 45  
 ctg ccg tgg gcc atc agc gcc ctg tta cta ccg aat ttc tgg gcg tat 192  
 Leu Pro Trp Ala Ile Ser Ala Leu Leu Leu Pro Asn Phe Trp Ala Tyr  
 50 55 60  
 tac agc tcg ttt agt ctc tct acg cga ctg cct gtt ttt ctg ttc ggc 240  
 Tyr Ser Ser Phe Ser Leu Ser Thr Arg Leu Pro Val Phe Leu Phe Gly  
 65 70 75 80  
 gct atg tgg ggg atc ggt aat atc aac tac ggc ctg acc atg cgt tat 288  
 Ala Met Trp Gly Ile Gly Asn Ile Asn Tyr Gly Leu Thr Met Arg Tyr  
 85 90 95

```

ctc ggc atg tcg atg gga att ggc atc gcc att ggc att acg ttg att 336
Leu Gly Met Ser Met Gly Ile Gly Ile Ala Ile Gly Ile Thr Leu Ile
100 105 110
gtc ggt acg ctg atg acg cca att atc aac ggc aat ttc gat gtg ttg 384
Val Gly Thr Leu Met Thr Pro Ile Ile Asn Gly Asn Phe Asp Val Leu
115 120 125
att agc acc gaa ggc gga cgc atg acg ttg ctc ggc gtt ctg gtg gcg 432
Ile Ser Thr Glu Gly Gly Arg Met Thr Leu Leu Gly Val Leu Val Ala
130 135 140
ctg att ggc gta ggg att gta act cgc gcc ggg cag ttg aaa gag cgc 480
Leu Ile Gly Val Gly Ile Val Thr Arg Ala Gly Gln Leu Lys Glu Arg
145 150 155 160
aag atg ggc att aaa gcc gaa gag ttc aat ctg aaa aaa ggg ctg gtg 528
Lys Met Gly Ile Lys Ala Glu Glu Phe Asn Leu Lys Lys Gly Leu Val
165 170 175
ctg gcg gtg atg tgc ggc att ttc tct gcc ggg atg tcc ttt gcg atg 576
Leu Ala Val Met Cys Gly Ile Phe Ser Ala Gly Met Ser Phe Ala Met
180 185 190
aac gcc gca aaa ccg atg cat gaa gcc gct gcc gca ctt ggc gtc gat 624
Asn Ala Ala Lys Pro Met His Glu Ala Ala Ala Leu Gly Val Asp
195 200 205
cca ctg tat gtc gct ctg cca agc tat gtt gtc atc atg ggc ggc ggc 672
Pro Leu Tyr Val Ala Leu Pro Ser Tyr Val Val Ile Met Gly Gly Gly
210 215 220
gcg atc att aac ctc ggt ttc tgt ttt att cgt ctg gca aaa gtg aag 720
Ala Ile Ile Asn Leu Gly Phe Cys Phe Ile Arg Leu Ala Lys Val Lys
225 230 235 240
gat ttg tcg cta aaa gcc gac ttc tcg ctg gca aaa tcg ctg atc att 768
Asp Leu Ser Leu Lys Ala Asp Phe Ser Leu Ala Lys Ser Leu Ile Ile
245 250 255
cac aat gtg tta ctc tcg aca ctg ggc ggg ttg atg tgg tat ctg caa 816
His Asn Val Leu Leu Ser Thr Leu Gly Gly Leu Met Trp Tyr Leu Gln
260 265 270
ttc ttt ttc tat gcc tgg ggc cac gcc cgc att ccg gcg cag tat gac 864
Phe Phe Phe Tyr Ala Trp Gly His Ala Arg Ile Pro Ala Gln Tyr Asp
275 280 285
tac atc agt tgg atg ctg cat atg agt ttc tat gta ttg tgc ggc ggt 912
Tyr Ile Ser Trp Met Leu His Met Ser Phe Tyr Val Leu Cys Gly Gly
290 295 300
atc gtc ggg ctg gtg ctg aaa gag tgg aac aat gca gga cgc cgt ccg 960
Ile Val Gly Leu Val Leu Lys Glu Trp Asn Asn Ala Gly Arg Arg Pro
305 310 315 320
gta acg gtg ttg agc ctc ggt tgt gtg gtg att att gtc gcc gct aac 1008
Val Thr Val Leu Ser Leu Gly Cys Val Val Ile Ile Val Ala Ala Asn
325 330 335
atc gtc ggc atc ggc atg gcg aat taa 1035
Ile Val Gly Ile Gly Met Ala Asn
340

```

&lt;210&gt; 19

&lt;211&gt; 344

&lt;212&gt; PRT

&lt;213&gt; Escherichia coli

&lt;400&gt; 19

Met Ser Asn Ala Ile Thr Met Gly Ile Phe Trp His Leu Ile Gly Ala

1

5

10

15

[illegible]